International Institute for Applied Systems Analysis
IIASA www.iiasa.ac.at

Confidence Regions for Linear Programs with Random Coefficients

Cipra, T.

IIASA Working Paper

WP-86-020

May 1986

NOT FOR QUOTATION
WITHOUT THE PERMISSION
OF THE AUTHOR

CONFIDENCE REGIONS FOR LINEAR PROGRAMS WITH RANDOM COEFFICIENTS

Tomáš Cipra

May 1986
WP-86-20

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
2361 Laxenburg, Austria

FOREWORD

Statistical approaches to stochastic optimization seem to be suitable from a practical point of view since observed data are frequently the only information we have on a stochastic optimization problem. In this paper, the possibility of constructing prediction regions in time series is utilized for probabilistic conclusions on the behavior of the solution of the problem.

The research was carried out within the Adaptation and Optimization Project of the System and Decision Sciences Program during the stay of the author as a guest scholar in IIASA.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

ACKNOWLEDGMENT

I would like to thank Roger Wets who discussed with me some problems from this work and gave me hints for further investigations.

AUTHOR

Assistant Professor Tomáš Cipra from the Faculty of Mathematics and Physics, Charles University, Prague works in the field of stochastic processes, stochastic programming and econometrics. He wrote this paper during his stay at IIASA in summer 1985.

ABSTRACT

If random values in a linear program with random coefficients can be predicted using previous observations on them one can utilize the appropriate prediction region and construct a confidence interval in which the optimal value of the objective function lies with a given probability (or even construct a confidence region for the optimal decision). It is a new statistical approach based on projection of the observed data into the time period of interest. The results are demonstrated by a numerical example.

CONFIDENCE REGIONS FOR LINEAR PROGRAMS WITH RANDOM COEFFICIENTS

Tomáš Cipra

1. INTRODUCTION

Let us have a linear program

$$\{\min c'x : Ax = b, x \geq 0\} \ , \tag{1.1}$$

where A, b, c, x have dimensions (m, n), $(m, 1)$, $(n, 1)$, $(n, 1)$, respectively. Let us consider the usual situation when A and c have known deterministic values and b is a random vector such that observations $y_1, \ldots, y_T$ of b in previous discrete time periods $t = 1, \ldots, T$ are at our disposal. If one succeeds in constructing a stochastic model generating the process $\{y_t\}$ one can also usually construct the prediction $\hat{y}_{T+h}$ of y_{T+h} for a time period $T + h$ in which the program (1.1) is to be solved (the most usual case is $h = 1$). Besides the point prediction $\hat{y}_{T+h}$ one can also construct a prediction region in which y_{T+h} lies with a prescribed probability (these prediction regions which are quite analogical to the confidence regions or intervals in theory of statistical estimation are even preferred for practical purposes in comparison with the point predictions).

There are various methods of predicting. If we confine ourselves to quantitative prediction methods only (so that the terms "projection" or "extrapolation" should be more suitable than "prediction" or "forecast") then the most important and usual representatives are the prediction method based on econometric modeling by means of systems of simultaneous equations (including the classical regression approach) and the prediction method in the framework of Box-Jenkins approach. Both methods are described briefly in section 2 of this paper.

If we have constructed the appropriate prediction region for b concerning the time period in which (1.1) is to be solved one can make use of it to obtain a confidence interval with a prescribed confidence probability for the optimal value of (1.1) or even (if performing more detailed analysis) to obtain a confidence region for the optimal decision in (1.1). It is obvious that linear parametric programming can be obtained in some way outside the model, see discussion in [3, p. 196] (e.g. in a simple regression situation it can be $x_{it} = t$ so that $\hat{x}_{i,T+h} = T + h$). Then under general assumptions on the stochastic behavior of the model in time $T + h$ (there must not be a change in the specification of the model in this time) the optimal point prediction for the endogenous variables in time $T + h$ can be constructed as

$$\hat{y}_{T+h} = \hat{\Pi}\hat{x}_{T+h} \tag{2.4}$$

with the $(1 - \alpha)100$ per cent prediction region of the form

$$(y_{T+h} - \hat{y}_{T+h})' S_{T+h,T+h}^{-1} (y_{T+h} - \hat{y}_{T+h}) \le \frac{(T-k)m}{T-k-m+1} F_{m,T-k-m+1}(\alpha) \; . \tag{2.5}$$

Here

$$S_{T+h,t+h} = \frac{1}{T-k}[1 + \hat{x}'_{T+h}\,(X'X)^{-1}\hat{x}_{T+h}](Y'Y - \hat{\Pi}X'Y) \tag{2.6}$$

is the estimated covariance matrix of the error $y_{T+h} - \hat{y}_{T+h}$ of the prediction and $F_{m,T-k-m+1}(\alpha)$ is the tabulated critical value of Fisher's distribution with the appropriate degrees of freedom and the level of significance α (e.g. (2.5) is the 95% prediction region for $\alpha = 0.05$). So called Hotelling's statistic has been used to derive (2.5).

REMARK 2 Various multivariate trends can be modeled by means of (2.1). E.g. the multivariate polynomial trend is modeled as

$$y_{it} = f_i(t) + v_{it},\ i = 1, \ldots, m,\ t = 1, \ldots, T \;, \tag{2.7}$$

where $f_i(t)$ is a polynomial of an order p_i (it means that the predetermined variables are chosen as powers of time t).

REMARK 3 Hymans [4] derived $(1 - \alpha)100$ per cent joint prediction intervals for particular components of y_{T+h} in the form

$$(\hat{y}_{i,T+h} - \sqrt{cs_{ii}},\ \hat{y}_{i,T+h} + \sqrt{cs_{ii}}),\ i = 1, \ldots, m \;, \tag{2.8}$$

where s_{ii} is the ith diagonal element of the matrix $(Y'Y - \hat{\Pi}X'Y)/(T - k)$ and

$$c = \frac{(T-k)m}{T-k-m+1}[1 + \hat{x}'_{T+h}\,(X'X)^{-1}\hat{x}_{T+h}]F_{m,T-k-m+1}(\alpha) \; . \tag{2.9}$$

positive–definite matrix and $k(\alpha)$ is a known constant. Let us denote this region as

$$P(\alpha) = \{b \in R^m : (b - \hat{b})' V(b - \hat{b}) \le k(\alpha)\} \ . \tag{2.16}$$

3. PROBLEM OF SOLVABILITY

Let us denote the region of solvability of the program (1.1) as

$$S = \{b \in R^m : (1.1) \text{ has an optimal solution } \} \tag{3.1}$$

(i.e. the program (1.1) is feasible and bounded for all $b \in S$) and assume that S is nonempty. Then in our context the problem of solvability consists in the investigation of the inclusion $P(\alpha) \subset S$.

Wets [11] deals with a general problem of this type when he investigates feasibility of stochastic programs with fixed recourse. Using theory of polar matrices and cone ordering he can treat cases with very general regions $P(\alpha)$. In our case we make use of the special elipsoid shape of the prediction region $P(\alpha)$ and proceed in the following way.

The solvability region S is a convex polyhedral cone with the vertex in the origin (see e.g. [9], [10]), i.e.

$$S = \{b \in R^m : h_i' b \ge 0, \ i = 1, \ldots, N\} \ . \tag{3.2}$$

The explicit numerical form of this cone (i.e. the vectors $h_1, \ldots, h_N$) can be found by means of various algorithmic procedures (e.g. [2], [7, p. 276]).

In order to simplify the solution of our problem let us transform the coordinate system in R^m so that the elipsoid (2.16) transfers to a sphere in R^m. The positive–definite matrix V from (2.16) can be decomposed as

$$V = C'C \ , \tag{3.3}$$

where C is an upper triangular matrix with positive elements on the main diagonal (so called Cholesky decomposition). If we define the transformation of the space R^m as

$$x \longrightarrow x^* = Cx, \ x \in R^m \tag{3.4}$$

(the asterisk will always denote the transformed value) then the elipsoid $P(\alpha)$ will be obviously transformed to the form

4. CONSTRUCTION OF CONFIDENCE REGIONS

Let us denote

$$\varphi(b) = \min\{c'x : Ax = b,\ x \geq 0\} \tag{4.1}$$

for $b \in S$. The function φ is convex, continuous and piecewise linear on S. More explicitly, there exist vectors $g_1, \ldots, g_r \in R^m$ such that

$$\varphi(b) = \max\{g_j'b : j = 1, \ldots, r\} \ . \tag{4.2}$$

According to the basis decomposition theorem (see [9]) the definition region S of the function φ can be decomposed to a finite number of convex polyhedral cones S_i with the vertices in the origin 0 such that the interiors of S_i are mutually disjunct and φ is linear on each of S_i (these regions correspond to particular bases B_i in A such that $\tilde{c}_i'\, B_i^{-1}\, A \leq c'$, where $\tilde{c}_i$ is the subvector of c corresponding to B_i). The explicit form of the cones S_i and the function φ can be also found by means of the mentioned algorithms [2] or [7].

Now let us try to determine the maximal and minimal values of φ over $P(\alpha)$. In other words, we shall construct the $(1 - \alpha)$ 100 per cent confidence interval for the optimal value of the objective function in (1.1).

THEOREM 2 *Let* $P(\alpha) \subset S$. *Then it holds*

$$\max\{\varphi(b) : b \in P(\alpha)\} = \max_{j=1,\ldots,r} \{\bar{g}_j'\,(\hat{b}^* + \frac{\sqrt{k(\alpha)}}{\|\bar{g}_j\|}\bar{g}_j)\} \ , \tag{4.3}$$

and

$$\min\{\varphi(b) : b \in P(\alpha)\} \geq \max_{j=1,\ldots,r} \{\bar{g}_j'\,(\hat{b}^* - \frac{\sqrt{k(\alpha)}}{\|\bar{g}_j\|}\bar{g}_j)\} \ , \tag{4.4}$$

where

$$\bar{g}_j = (C^{-1})'g_j,\ j = 1, \ldots, r \tag{4.5}$$

and for $g_j = \bar{g}_j = 0$ *we put*

$$\bar{g}_j'(\hat{b}^* \pm \frac{\sqrt{k(\alpha)}}{\|\bar{g}_j\|}\bar{g}_j) = 0 \ .$$

PROOF We can write

$$\max\{\varphi(b) : b \in P(\alpha)\} = \max_{b \in P(\alpha)} \{\max_{j=1,\ldots,r} \{g_j'\, b\}\}$$

REMARK 7 This work is not the first one dealing with confidence regions in linear programs with random coefficients. E.g., results have been obtained by means of projection of rectangulars in which the values (A, B, c) lie with a given probability (see [6, section 13.1] or [8]). These rectangulars are defined by means of the mean values and standard deviations of the random components of (A, b, c) and do not make use of the correlation structure (relations among particular random components) how it is the case when projecting elipsoids.

If we carry out a more detailed analysis of (1.1) as the problem of parametric programming with the (vector) parameter b (i.e. if one finds explicitly the decomposition of the solvability region S to the cones S_i) then it is even possible to construct a confidence region for the optimal decision in (1.1). We shall show such construction including the application of Theorem 1 and Corollary of Theorem 2 in the following example which is simple enough to demonstrate clearly the previous theory. The application for real examples assumes the exploitation of software from statistics and linear parametric programming.

EXAMPLE The authors of [7] investigated the following problem

$$\begin{array}{ll} & \min\{x_2 + x_3 + 3x_4\} \\ s.t. & x_1 - 2x_2 + x_3 - x_4 + x_5 = b_1 \\ & 2x_1 + 3x_2 - x_3 + 2x_4 + x_6 = b_2 \\ & -x_1 + 2x_2 + 3x_3 - 3x_4 + x_7 = b_3 \\ & x_1, \ldots, x_7 \geq 0 \end{array} \tag{4.11}$$

In this case it is

$$S = \{b \in R^3 : 3b_1 + 2b_2 \geq 0,\ 3b_2 + b_3 \geq 0,\ b_1 + b_2 \geq 0,\ 13b_1 + 8b_2 + b_3 \geq 0,\ 3b_2 + 2b_3 \geq 0\} \tag{4.12}$$

and

$$\varphi(b) = \max\{0,\ -\frac{1}{2}b_1,\ -b_2,\ -\frac{3}{4}b_1 - \frac{3}{4}b_3\ ,\ -\frac{9}{8}b_1 - \frac{5}{8}b_3,\ -\frac{3}{4}b_2 - \frac{3}{2}b_3\} \tag{4.13}$$

for $b \in S$. Table 1 contains the description of all cones S_i from the decomposition of S including the corresponding forms of φ and the optimal bases B_i.

Let the $(1-\alpha)100$ per cent prediction region (2.16) have the following form

$$\begin{bmatrix} b_1 + 15.5 \\ b_2 - 142.7 \\ b_3 - 87.3 \end{bmatrix}' \begin{bmatrix} 6.25 & -3 & 0.75 \\ -3 & 19.93 & -9.39 \\ 0.75 & -9.39 & 17.46 \end{bmatrix} \begin{bmatrix} b_1 + 15.5 \\ b_2 - 142.7 \\ b_3 - 87.3 \end{bmatrix} \leq 6.76\ , \tag{4.14}$$

According to (4.12) the vectors h_i from (3.2) are

$$h_1 = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}, \quad h_2 = \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}, \quad h_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad h_4 = \begin{bmatrix} 13 \\ 8 \\ 1 \end{bmatrix}, \quad h_5 = \begin{bmatrix} 0 \\ 3 \\ 2 \end{bmatrix} \tag{4.16}$$

and according to (4.13) the vectors g_j from (4.2) are

$$\begin{gathered} g_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad g_2 = \begin{bmatrix} -1/2 \\ 0 \\ 0 \end{bmatrix}, \quad g_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}, \quad g_4 = \begin{bmatrix} -3/4 \\ 0 \\ -3/4 \end{bmatrix}, \\ g_5 = \begin{bmatrix} -9/8 \\ 0 \\ -5/8 \end{bmatrix}, \quad g_6 = \begin{bmatrix} 0 \\ -3/4 \\ -3/2 \end{bmatrix} \end{gathered} \tag{4.17}$$

The left-hand-sides of the inequalities (3.8) (or equivalently of (3.9) or (3.10)) are

$$235.03,\ 512.86,\ 125.76,\ 1010.72,\ 599.61\ . \tag{4.18}$$

Since each of these values is non-negative the elipsoid (4.14) is the subset of the solvability region (4.12). The corresponding confidence region for the optimal value of the objective function is according to (4.9) or (4.10)

$$[7.21,\ 8.29]\ . \tag{4.19}$$

If the vector $\hat{b}$ is replaced by the vector

$$\hat{b} = (43.5,\ 112.8,\ 90.3)' \tag{4.20}$$

then the problem stays solvable and the confidence interval is

$$[0,\ 0] \tag{4.21}$$

so that the objective function has the optimal value 0 with the probability at least $1 - \alpha$.

The cones S_i can be written formally as

$$S_i = \{b \in R^3 : H_i b \geq 0\},\ i = 1, \ldots, 9\ , \tag{4.22}$$

where H_i are (3,3) matrices, e.g. it is

$$H_3 = \begin{bmatrix} 0 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$$

Then it is not difficult to verify that the $(1 - \alpha)100$ per cent confidence region for the optimal decision in (4.11) can be taken as

5. APPLICATION OF BUNCHING METHOD

The method of bunching [12] and especially its trickling down modification in combination with Schur-complement bases updates [13] is the efficient tool for solving linear programs with variable right-hand-sides (a bunch is such subset of a given set of right-hand-side vectors which corresponds to the same optimal basis of the program).

In our case the bunching method will enable to solve in an efficient way a lot of problems of the type

$$\{\min c'x : Ax = z^k,\ x \geq 0\},\ k = 1, \ldots, K \tag{5.1}$$

(the points z^k are chosen from the elipsoid $P(\alpha)$) without performing explicitly the decomposition of the solvability region S to the cones S_i.

As the choice of the points z^k is concerned one can use various strategies. E.g., it is possible to choose the points z^k randomly from the surface of the elipsoid $P(\alpha)$. If we transform the coordinates according to (3.4) then we can generate these points uniformly from the surface of the sphere (3.5) taking

$$\begin{aligned}
z_1^* &= \hat{b}_1^* + \sqrt{k(\alpha)} \cos \vartheta_1 \cos \vartheta_2 \cos \vartheta_3 \cdots \cos \vartheta_{m-1} \\
z_2^* &= \hat{b}_2^* + \sqrt{k(\alpha)} \sin \vartheta_1 \cos \vartheta_2 \cos \vartheta_3 \cdots \cos \vartheta_{m-1} \\
z_3^* &= \hat{b}_3^* + \sqrt{k(\alpha)} \sin \vartheta_2 \cos \vartheta_3 \cos \vartheta_4 \cdots \cos \vartheta_{m-1} \\
z_4^* &= \hat{b}_4^* + \sqrt{k(\alpha)} \sin \vartheta_3 \cos \vartheta_4 \cos \vartheta_5 \cdots \cos \vartheta_{m-1} \\
&\vdots \\
z_{m-1}^* &= \hat{b}_{m-1}^* + \sqrt{k(\alpha)} \sin \vartheta_{m-2} \cos \vartheta_{m-1} \\
z_m^* &= \hat{b}_m^* + \sqrt{k(\alpha)} \sin \vartheta_{m-1} ,
\end{aligned} \tag{5.2}$$

where $0 \leq \vartheta_1 \leq 2\pi$, $-\pi/2 \leq \vartheta_2 \leq \pi/2, \ldots, -\pi/2 \leq \vartheta_{m-1} \leq \pi/2$ are independent random variables with uniform distributions on their ranges.

The trickling down procedure can be started in the point $\hat{b}^*$ (the center of the sphere $P^*(\alpha)$). Let $B_{(1)}^*$ be the optimal basis for this point $\hat{b}^*$ (it holds obviously $B_{(1)}^* = CB_{(1)}$, where $B_{(1)}$ is the optimal basis for the point $\hat{b}$ before the transformation (3.4) since the problem (1.1) can be written equivalently as $\{\min c'x : CAx = Cb,\ x \geq 0\}$). Let $z^{1*} = (z_1^{1*}, \ldots, z_m^{1*})'$ be the first point generated according to (5.2). By using trickling down procedure (i.e. the proper se-

REFERENCES

1 Box, G.E.P., Jenkins, G.M. Time Series Analysis, Forecasting and Control. Holden Day, San Francisco 1970.

2 Gal, T., Nedoma, J. Multiparametric linear programming. Management Science 18, 1972, 406–422.

3 Granger, C.W.J., Newbold, P. Forecasting Economic Time Series. Academic Press, New York 1977.

4 Hymans, S.H. Simultaneous confidence intervals in econometric forecasting. Econometrica 36, 1968, 18–30.

5 Johnston, J. Econometric Methods. McGraw-Hill, New York 1972.

6 Judin, D.B. Mathematical Methods of Control under Incomplete Information (Problems and Methods of Stochastic Programming). Sov. radio, Moskva 1974 (in Russian).

7 Nožička, F., Guddat, J., Hollatz, H., Bank, B. Theorie der linearen parametrischen Optimierung. Akademie-Verlag, Berlin 1974.

8 Theodorescu, R. Random programs. Math. Operationsforsch. Statist. 3, 1972, 19–47.

9 Walkup, D., Wets, R. Lifting projections of convex polyedra. Pacific J. Mathem. 28, 1969, 465–475.

10 Wets, R. Programming under uncertainty: the equivalent convex program. J. SIAM Appl. Math. 14, 1966, 89–105.

11 Wets, R. Stochastic programs with fixed recourse: the equivalent deterministic program. SIAM Review 16, 1974, 309–339.

12 Wets, R. Stochastic programming: solution techniques and approximation schemes. In: Mathematical Programming: The State of the Art (A. Bachem, M. Grotsched, B. Korte eds.). Springer, Berlin 1983, 566–603.

13 Wets, R. Large scale linear programming techniques in stochastic programming. IIASA Working Paper, WP-84-90, Laxenburg, Austria 1984.

may be convenient tool for this purpose. In the initial stage of such analysis based on prediction regions one should investigate whether the prediction region for b with a prescribed confidence probability (e.g. 95%) is the subset of such region for b in which the problem (1.1) is solvable (so called solvability region). This problem is discussed in section 3 while the construction of the confidence regions is described in section 4 including a simple example which demonstrates it. A possible application of so called bunching method [12], [13] for the considered situation is suggested in section 5.

REMARK 1 Although the simplest case with random b only is considered in this work the method could be extended in principal to more general situations. The case with random c only is equivalent to the case discussed here due to duality.

2. CONSTRUCTION OF PREDICTION REGIONS IN PRACTICE

In this section the both mentioned methods of quantitative prediction are reminded:

a) Prediction based on econometric modeling (see e.g. [5]) is a general method which includes as a special case e.g. the prediction based on the classical regression analysis. If using this method we must e.g. have at our disposal the estimated model of simultaneous equations in the reduced form

$$y_t = \Pi x_t + v_t,\ t = 1, \ldots, T\ . \tag{2.1}$$

Here $y_t = (y_{1t}, \ldots, y_{mt})'$ is a vector of endogenous variables in time t which is explained by a vector of predetermined variables $x_t = (x_{1t}, \ldots, x_{kt})'$ in time t (the object of the prediction are the endogenous variables), Π is a (m, k) matrix of parameters and $v_t = (v_{1t}, \ldots, v_{mt})'$ is a vector of disturbances in time t. One assumes that $Ev_t = 0$, $Ev_t v_t' = \sum_{vv}$ (a positive-definite matrix) and $Ev_s v_t' = 0$ for $s \neq t$. The model can be summarized for all t as

$$Y = X\Pi' + V\ , \tag{2.2}$$

where $Y = (y_1, \ldots, y_T)'$, $X = (x_1, \ldots, x_T)'$ and $V = (v_1, \ldots, v_T)'$. The classical OLS (Ordinary Least Squares) estimator of Π has the form

$$\hat{\Pi} = Y'X(X'X)^{-1}\ . \tag{2.3}$$

Let $\hat{x}_{T+h}$ be a vector of predicted predetermined variables for time $T + h$ which

b) Prediction in the framework of Box-Jenkins approach (see [1], [3]) is exploited by many statisticians and econometricians as a very flexible and fruitful prediction method. Similarly as in the econometric modeling one must construct an appropriate model at first. Box-Jenkins methodology utilizes so called ARMA (p, q) models (or their various modifications) of the form

$$y_t + A_1 y_{t-1} + \cdots + A_p y_{t-p} = \varepsilon_t + B_1 \varepsilon_{t-1} + \cdots + B_q \varepsilon_{t-q} \ , \tag{2.10}$$

where y_t is the modeled m-dimensional process, $A_1, \ldots, A_p$ and $B_1, \ldots, B_q$ are (m, m) matrices of parameters and ε_t is the m-dimensional white noise, i.e. $E\varepsilon_t = 0$, $E\varepsilon_t \varepsilon_t' = \sum$ (a positive-definite matrix), $E\varepsilon_s \varepsilon_t' = 0$ for $s \neq t$. Then under general conditions the optimal point prediction for time $T + h$ can be written as

$$\hat{y}_{T+h} = \sum_{j=0}^{\infty} C_{j+h} \varepsilon_{T-j} \ , \tag{2.11}$$

where the (m, m) matrices C_i fulfill the following power series equation

$$\begin{aligned}(I + A_1 z + \cdots + A_p z^p)(I + C_1 z + C_2 z^2 + \cdots) \\ = I + B_1 z + \cdots + B_q z^q \ .\end{aligned} \tag{2.12}$$

The corresponding $(1 - \alpha)100$ per cent prediction region has the form

$$(y_{T+h} - \hat{y}_{T+h})' V(h)^{-1} (y_{T+h} - \hat{y}_{T+h}) \leq \chi_m^2(\alpha) \ , \tag{2.13}$$

where

$$V(h) = \sum_{j=0}^{h-1} C_j \sum C_j' \tag{2.14}$$

is the covariance matrix of the error $y_{T+h} - \hat{y}_{T+h}$ of the prediction and $\chi_m^2(\alpha)$ is the tabulated critical value of chi-squared distribution with m degrees of freedom and the level of significance α.

The following conclusion can be drawn from the previous text. In both prediction methods (and also in other less important ones) the appropriate $(1 - \alpha)100$ per cent prediction region has the geometric form of an elipsoid. This elipsoid can be written generally for the program (1.1) as

$$(b - \hat{b})' V (b - \hat{b}) \leq k(\alpha) \ , \tag{2.15}$$

where $\hat{b}$ is a known vector (the center of the elipsoid), V is a known

$$P^*(\alpha) = \{b^* \in R^m : (b^* - \hat{b}^*)'(b^* - \hat{b}^*) \leq k(\alpha)\} \ , \tag{3.5}$$

which is a m-dimensional sphere with the center $\hat{b}^*$ and the radius $\sqrt{k(\alpha)}$. The transformed solvability region S has the form

$$S^* = \{b^* \in R^m : \bar{h}_i' b^* \geq 0, \ i = 1, \ldots, N\} \ , \tag{3.6}$$

where

$$\bar{h}_i = (C^{-1})' h_i, \ i = 1, \ldots, N \ . \tag{3.7}$$

If the interior of the elipsoid $P(\alpha)$ (denoted as int $P(\alpha)$) contains the zero vector 0 (or equivalently if the zero vector 0 lies in int $P^*(\alpha)$) then the problem discussed in this section has the following simple solution.

LEMMA 1 *Let* $0 \in int P(\alpha)$. *Then the inclusion* $P(\alpha) \subset S$ *is true if and only if* $S = R^m$.

PROOF Lemma is obvious since S is a cone with the vertex in 0 and $P(\alpha)$ is an elipsoid.

General solution of the considered problem is given in the following theorem.

THEOREM 1 *The inclusion* $P(\alpha) \subset S$ *is true if and only if it holds*

$$\bar{h}_i' (\hat{b}^* - \frac{\sqrt{k(\alpha)}}{\|\bar{h}_i\|} \bar{h}_i) \geq 0, \ i = 1, \ldots, N \ , \tag{3.8}$$

where $\| \ \|$ *is the usual Eucleidian norm in* R^m.

PROOF $P(\alpha) \subset S$ is equivalent to $P^*(\alpha) \subset S^*$ and this last inclusion holds iff the sphere $P^*(\alpha)$ with the center $\hat{b}^*$ and the radius $\sqrt{k(\alpha)}$ lies in all half-spaces $\{b^* \in R^m : \bar{h}_i' b^* \geq 0\}$, $i = 1, \ldots, N$. This last condition is obviously equivalent to (3.8).

REMARK 4 The inequalities (3.8) can be written in the equivalent form

$$h_i' \hat{b} - \sqrt{k(\alpha)} \|\bar{h}_i\| \geq 0, \ i = 1, \ldots, N \tag{3.9}$$

or

$$h_i' \hat{b} - \sqrt{k(\alpha) h_i' V^{-1} h_i} \geq 0, \ i = 1, \ldots, N \ . \tag{3.10}$$

$$= \max_{b^* \in P^*(\alpha)} \{ \max_{j=1,\ldots,r} \{\bar{g}_j' b^*\}\} = \max_{j=1,\ldots,r} \{ \max_{b^* \in P^*(\alpha)} \{\bar{g}_j' b^*\}\} .$$

Now the relation (4.3) is proved since it obviously holds

$$\max_{b^* \in P^*(\alpha)} \{\bar{g}_j' b^*\} = \bar{g}_j' (\hat{b}^* + \frac{\sqrt{k(\alpha)}}{\|\bar{g}_j\|}\bar{g}_j) \tag{4.6}$$

(we maximize the linear function $\bar{g}_j' b^*$ over the sphere with the center $\hat{b}^*$ and the radius $\sqrt{k(\alpha)}$; the maximal value is achieved in the point where the vector directed from $\hat{b}^*$ as the gradient $\bar{g}_j$ of the function $\bar{g}_j' b^*$ crosses the surface of the sphere, i.e. in the point $\hat{b}^* + (\sqrt{k(\alpha)}/\|\bar{g}_j\|)\bar{g}_j$).

As the relation (4.4) is concerned we have

$$\min\{\varphi(b) : b \in P(\alpha)\} = \min_{b^* \in P^*(\alpha)} \{ \max_{j=1,\ldots,r} \{\bar{g}_j' b^*\}\} \geq \max_{j=1,\ldots,r} \{ \min_{b^* \in P^*(\alpha)} \bar{g}_j' b^*\}\} .$$

The last inequality holds since it is

$$\max_{j=1,\ldots,r} \{\bar{g}_j' b^*\} \geq \max_{j=1,\ldots,r} \{ \min_{b^* \in P^*(\alpha)} \{\bar{g}_j' b^*\}\} \tag{4.7}$$

for each $b^* \in P^*(\alpha)$ so that we can replace the left-hand-side of (4.7) by its minimal value over $b^* \in P^*(\alpha)$. The proof is finished because one can derive in the same way as (4.6) that

$$\min_{b^* \in P^*(\alpha)} \{\bar{g}_j' b^*\} = \bar{g}_j' (\hat{b}^* - \frac{\sqrt{k(\alpha)}}{\|\bar{g}_j\|}\bar{g}_j) . \tag{4.8}$$

COROLLARY *Let $P(\alpha) \subset S$. Then the interval of the form*

$$[\max_{j=1,\ldots,r} \{\bar{g}_j' (\hat{b}^* - \frac{\sqrt{k(\alpha)}}{\|\bar{g}_j\|}\bar{g}_j)\}, \max_{j=1,\ldots,r} \{\bar{g}_j' (\hat{b}^* + \frac{\sqrt{k(\alpha)}}{\|\bar{g}_j\|}\bar{g}_j)\}] \tag{4.9}$$

is the confidence interval for the optimal value of the objective function in (1.1) with the confidence probability at least $1 - \alpha$.

REMARK 5 The interval (4.9) can be written again in the equivalent form

$$[\max_{j=1,\ldots,r} \{g_j' \hat{b} - \sqrt{k(\alpha)}\|\bar{g}_j\|\}, \max_{j=1,\ldots,r} \{g_j' \hat{b} + \sqrt{k(\alpha)}\|\bar{g}_j\|\}] . \tag{4.10}$$

REMARK 6 Since the function $\varphi(b)$ attains the value $+\infty$ outside the set S (see e.g. [11]) one can omit the assumption $P(\alpha) \subset S$ in the previous Corollary and formulate it in such a way that the optimal objective value lies in the interval (4.9) or is equal to $+\infty$ with the probability at least $1 - \alpha$.

Table 1 The analysis of the linear parametric problem (4.11).

i	S_i	$\varphi(b)$	B_i (numbers of columns of A)
1	$b_1 \geq 0,\ b_2 \geq 0,\ b_3 \geq 0$	0	(5, 6, 7)
2	$b_1 \geq 0,\ -2b_1 + b_2 \geq 0,$ $b_1 + b_3 \geq 0$	0	(1, 6, 7)
3	$-b_3 \geq 0,\ b_1 + b_3 \geq 0,$ $b_2 + 2b_3 \geq 0$	0	(1, 5, 6)
4	$b_1 \geq 0,\ 2b_1 - b_2 \geq 0,$ $b_2 + 2b_3 \geq 0$	0	(1, 5, 7)
5	$-b_1 \geq 0,\ 3b_1 + 2b_2 \geq 0,$ $b_1 + b_3 \geq 0$	$-\frac{1}{2}b_1$	(2, 6, 7)
6	$-b_2 \geq 0,\ b_1 + b_2 \geq 0,$ $3b_2 + b_3 \geq 0$	$-b_2$	(3, 5, 7)
7	$-b_1 - b_3 \geq 0,\ 3b_1 - b_3 \geq 0,$ $-b_1 + b_2 + b_3 \geq 0$	$-\frac{3}{4}b_1 - \frac{3}{4}b_3$	(1, 4, 6)
8	$-b_1 - b_3 \geq 0,\ -3b_1 + b_3 \geq 0$ $13b_1 + 8b_2 + b_3 \geq 0$	$-\frac{9}{8}b_1 - \frac{5}{8}b_3$	(2, 4, 6)
9	$-b_2 - 2b_3 \geq 0,\ 3b_2 + 2b_3 \geq 0,$ $b_1 - b_2 - b_3 \geq 0$	$-\frac{3}{4}b_2 - \frac{3}{2}b_3$	(1, 4, 5)

i.e.

$$\hat{b} = \begin{bmatrix} -15.5 \\ 142.7 \\ 87.3 \end{bmatrix}, \quad V = \begin{bmatrix} 6.25 & -3 & 0.75 \\ -3 & 19.93 & -9.39 \\ 0.75 & -9.39 & 17.46 \end{bmatrix}, \quad C = \begin{bmatrix} 2.5 & -1.2 & 0.3 \\ 0 & 4.3 & -2.1 \\ 0 & 0 & 3.6 \end{bmatrix} \tag{4.15}$$

$$\bigcup_{i=1}^{9} \{x \in R^7 : (\tilde{x}_i - B_i^{-1}\hat{b})' B_i' V B_i (\tilde{x}_i - B_i^{-1}\hat{b}) \le k(\alpha),\ H_i B_i \tilde{x}_i \ge 0\} \quad , \tag{4.23}$$

where $\tilde{x}_i$ denotes such subvector of the vector $x \in R^7$ which corresponds to the basis B_i. E.g. if we use the prediction region (4.14) then the set with the index $i = 1$ in the union (4.23) has the form

$$\{x \in R^7 : \begin{bmatrix} x_5 + 15.5 \\ x_6 - 142.7 \\ x_7 - 87.3 \end{bmatrix}' \begin{bmatrix} 6.25 & -3 & 0.75 \\ -3 & 19.93 & -9.39 \\ 0.75 & -9.39 & 17.46 \end{bmatrix} \begin{bmatrix} x_5 + 15.5 \\ x_6 - 142.7 \\ x_7 - 87.3 \end{bmatrix} \le 6.76 \quad , \tag{4.24}$$

$$x_1 = x_2 = x_3 = x_4 = 0,\ x_5 \ge 0,\ x_6 \ge 0,\ x_7 \ge 0\}$$

and the set with the index $i = 2$ has the form

$$\{x \in R^7 : \begin{bmatrix} x_1 + 15.5 \\ x_6 - 173.7 \\ x_7 - 71.8 \end{bmatrix}' \begin{bmatrix} 127.49 & 46.25 & -35.49 \\ 46.25 & 19.93 & -9.39 \\ -35.49 & -9.39 & 17.46 \end{bmatrix} \begin{bmatrix} x_1 + 15.5 \\ x_6 - 173.7 \\ x_7 - 71.8 \end{bmatrix} \le 6.76 \quad , \tag{4.25}$$

$$x_2 = x_3 = x_4 = x_5 = 0,\ x_1 \ge 0,\ x_6 \ge 0,\ x_7 \ge 0\} \ .$$

It is obvious that the described method of the construction of the confidence regions exploits substantially the procedures of linear parametric programming so that its practical applicability is limited if the dimensions of the problem are large (one must also keep in mind the fact that for increasing m the construction of the prediction elipsoid becomes more and more difficult). The method seems to be suitable in such cases when one solves a lot of problems (1.1) with the same values A and c for various trajectories $\{y_t\}$ so that the tedious and expensive procedures of parametric programming will bring an effect (such situations may be usual in routine practical problems). On the other hand, the method cannot be recommended for a single use in large scale problems for which more effective procedures should be suggested. One of such potential procedures based on the bunching method is sketched in the following section.

quence of dual simplex steps exploiting Schur-complement updates) one will find the corresponding sequence $B^*_{(1)}, \ldots, B^*_{(s)}$ of the bases which is ended by the basis $B^*_{(s)}$ optimal for the point z^{1*}. Let us calculate the values

$$l^1 = \bar{g}'_{(s)}\left[\hat{b}^* - \frac{\sqrt{k(\alpha)}}{\|\bar{g}_{(s)}\|}\bar{g}_{(s)}\right], \quad u^1 = \bar{g}'_{(s)}\left[\hat{b}^* + \frac{\sqrt{k(\alpha)}}{\|\bar{g}_{(s)}\|}\bar{g}_{(s)}\right], \tag{5.3}$$

where

$$\bar{g}_{(s)} = (B^{*-1}_{(s)})'\tilde{c}_{(s)} \tag{5.4}$$

($\tilde{c}_{(s)}$ is the subvector of c corresponding to the basis $B^*_{(s)}$). The same procedure will be performed with the second generated point z^{2*} producing the values l^2 and u^2, etc. If proceeding in this way we obtain a tree rooted at the basis $B^*_{(1)}$ (see [13]) the paths of which are ended by the couples (l^1, u^1), (l^2, u^2), ..., (l^K, u^K). The $(1-\alpha)100$ per cent prediction interval for the optimal objective value can be then approximated by the interval

$$\left[\max_{k=1,\ldots,K} l^{(k)}, \max_{k=1,\ldots,K} u^{(k)}\right]. \tag{5.5}$$

The stopping rule by means of which the number K is found can be prescribed in such a way that the last L couples $(l^{K-L+1}, u^{K-L+1}), \ldots, (l^K, u^K)$ will satisfy

$$\begin{aligned} \max_{k=1,\ldots,K} l^k - \max_{k=1,\ldots,L-K} l^k < \varepsilon, \\ \max_{k=1,\ldots,K} u^k - \max_{k=1,\ldots,L-K} u^k < \varepsilon, \end{aligned} \tag{5.6}$$

where an integer L and a sufficiently small $\varepsilon > 0$ are chosen apriori.

More complicated strategies can be suggested but the previous one seems to be suitable in spite of its simplicity. The inaccuracies which can originate when using the generating formulas (5.2) do not reduce the efficiency of the method since the points $z^{1*}, z^{2*}, \ldots$ are used only to determine the corresponding optimal bases and these bases do not usually vary in the neighborhoods of particular right-hand-side vectors. Moreover, when the components of the vector $\hat{b}$ are large (as it is frequent in practice) then usually only small number of the cones S_l from the decomposition of S have nonempty intersections with the elipsoid $P(\alpha)$ so that the mentioned tree from the trickling down procedure has small number of paths which reduces the computing effort.